R E P O R T   R E S U M E S

ED 018 831   24   CG 001 766

INTRA-INDIVIDUAL VARIABILITY OF BEHAVIOR AND THE PREDICTABILITY OF ACADEMIC SUCCESS.

BY- BERDIE, RALPH F.
MINNESOTA UNIV., MINNEAPOLIS

REPORT NUMBER BR-6-8694   PUB DATE FEB 68
CONTRACT OEC-3-7-068694-2082
EDRS PRICE MF-~~$0.25~~ $0.50 HC-$2.04 ~~40P.~~ 60P.

DESCRIPTORS- PREDICTION, ENGINEERING EDUCATION, *ACADEMIC PERFORMANCE, ACADEMIC ACHIEVEMENT, *PSYCHOLOGICAL TESTING, BEHAVIOR PATTERNS, *BEHAVIOR CHANGE, *TEST RELIABILITY,

THE PURPOSE OF THE RESEARCH WAS TO OBSERVE THE CONSISTENCY OF INTRA-INDIVIDUAL VARIABILITY FROM ONE TASK TO ANOTHER, TO DETERMINE THE RELIABILITY OF OBSERVATIONS OF INTRA-INDIVIDUAL VARIABILITY, AND TO OBSERVE THE EXTENT TO WHICH THE ACCURACY OF PREDICTION VARIED WITH THE CONSISTENCY OF THE INDIVIDUAL'S BEHAVIOR. MALE ENGINEERING STUDENTS AT THE UNIVERSITY OF MINNESOTA WERE ADMINISTERED A BATTERY OF SIX BRIEF, INDEPENDENT TESTS ON EACH OF TWENTY DAYS. MOST OF THE OBSERVATIONS OF INTRA-INDIVIDUAL VARIABILITY WERE RELIABLE. VARIABILITY OVER TIME ON SOME TASKS IS RELATED TO VARIABILITY ON OTHER TASKS, BUT THESE RELATIONSHIPS ARE ONLY MODERATE AND THE HIGHEST CORRELATIONS BETWEEN ANY TWO OF THE VARIANCE INDICES WAS ONLY ABOUT .50. INTRA-INDIVIDUAL VARIABILITY IS NOT SPECIFIC TO EACH TASK BUT NEITHER IS THERE A BROADLY GENERALIZED CHARACTERISTIC OF SUCH VARIABILITY. A RELATIONSHIP WAS OBSERVED BETWEEN THE VARIABILITY OVER TIME OF A PERSON'S BEHAVIOR ON SIMPLE TASKS AND THE EFFECTIVENESS WITH WHICH HIS ACADEMIC PERFORMANCE CAN BE PREDICTED. SOME OF THE ANALYSES REVEALED ERRORS OF PREDICTION SIGNIFICANTLY SMALLER FOR THE LESS VARIABLE GROUP. CONSIDERED IN THE LIGHT OF RESULTS OF OTHER RESEARCH, THESE RESULTS SUGGEST THAT A COMPLEX BUT PRESUMEDLY USEFUL RELATIONSHIP EXITS BETWEEN THE CONSISTENCY OF AN INDIVIDUAL'S BEHAVIOR AND THE EXTENT TO WHICH HIS ACADEMIC BEHAVIOR IS PREDICTABLE. (AUTHOR)

CG
from CSRC

FINAL REPORT
Project No. 6-8694 -24
Contract No. OEC-3-7-068694-2082

INTRA-INDIVIDUAL VARIABILITY OF BEHAVIOR AND THE PREDICTABILITY OF ACADEMIC SUCCESS

February 1968

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

OE 6000 (REV. 9-66)

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

OFFICE OF EDUCATION

ERIC REPORT RESUME

ERIC ACCESSION NO.				
CLEARINGHOUSE ACCESSION NUMBER	RESUME DATE 02-15-68	P.A.	T.A.	IS DOCUMENT COPYRIGHTED? YES ☐ NO ☒ ERIC REPRODUCTION RELEASE? YES ☒ NO ☐

TITLE

Intra-individual Variability of Behavior and the Predictability of Academic Success. Final Report.

PERSONAL AUTHOR(S)

Berdie, Ralph F.

INSTITUTION (SOURCE)

University of Minnesota, Minneapolis, Minnesota, Office of the Dean of Students

SOURCE CODE

REPORT/SERIES NO. Project No. 6-8694, Contract No. OEC-3-7-068694-2082

OTHER SOURCE

U.S. Office of Education

SOURCE CODE

OTHER REPORT NO. Project No. 6-8694, Contract No. OEC-3-7-068694-2082

OTHER SOURCE

SOURCE CODE

OTHER REPORT NO.

PUB'L. DATE 2-25-68

CONTRACT/GRANT NUMBER OEC-3-7-068694-2082

PAGINATION, ETC.

59

RETRIEVAL TERMS

Intra-individual variability, variability, prediction, moderater variable, repeated psychological measurement, engineering education, consistency of behavior, behavior consistency.

IDENTIFIERS

None

ABSTRACT

The purpose of the research was to observe the consistency of intra-individual variability from one task to another, to determine the reliability of observations of intra-individual variability, and to observe the extent to which the accuracy of prediction varied with the consistency of the individual's behavior. One hundred male engineering students at the University of Minnesota were administered a battery of six brief independent tests, the Repeated Psychological Measurements, on each of twenty days, a different form of each test being used each day. For each subject variance indices were calculated on each of the six tests and on the total of seven tests and the relationships between variance indices were observed, the reliabilities of the indices determined, and the relationships between intra-individual variability and the prediction of academic success observed.

Most of the observations of intra-individual variability were reliable. Variability over timeon some tasks is related to variability on other tasks, but these relationships areonly moderate and the highest correlations between any two of the variance indices was only about .50. Intra-individual variability is not specific to each task but neither is there a broadly generalized characteristic of such variability. A relationship was observed between the variability over time of a person's behavior on simple tasks and the effectiveness with which his academic performance can be predicted. Some of the analyses revealed errors of prediction significantly smaller for the less variable group. One analysis showed that the correlation between predicted grades and obtained grades was .67 for a less variable group, as compared to .16 for a more variable group. Considered in the light of results of other research, these results suggest that a complex but presumedly useful relationship exists between the consistency of an individual's behavior and the extent to which his academic behavior is predictable.

INTRA-INDIVIDUAL VARIABILITY OF BEHAVIOR AND THE PREDICTABILITY OF ACADEMIC SUCCESS

Project No. 6-8694
Contract No. OEC-3-7-068694-2082

Ralph F. Berdie

February 1968

The research reported herein was performed pursuant to a contract with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

University of Minnesota

Minneapolis, Minnesota

The Consistency and Generalizability of Intra-individual Variability*

by

Ralph F. Berdie

Office of the Dean of Students

University of Minnesota

The extent to which an individual's behavior is consistently variable on a given task or on different tasks may provide cues as to the effectiveness of his behavior. The variability studied here corresponds to that described by Fiske (Fiske & Maddi, p.327), ". . . one form of variability, the variation in the behavior of a given organism at different times but under the same external conditions." This form of variability is to be contrasted with that discussed by Hull (1927) who was concerned with variability in the amount of different traits possessed by an individual, and also to be contrasted with the variability discussed by Wechsler (1950) who was concerned with the variability of a given trait within the population.

Theoretically, an individual cannot reproduce a behavior identically insofar as once he has performed a task, repetition of that task must be influenced by its prior performance. In spite of the impossibility of studying the individual's variability while performing the same task,

---

*This project was supported with funds from the U.S.O.E.-O.E.C.-3-7-068694-2082. Appreciation is expressed to Mr. Richard Arvey and Miss Diane Johnson for their contribution in analyzing the data.

his variability can be studied while performing a homogenous group of tasks. Tasks and situations change from performance to performance, but these can be ordered into highly similar categories and variability of behavior so studied.

Fiske provided one theoretical basis for the analysis of variability. He regarded individual variability as having a coping function which effects the organism's adaptability. When several alternate behaviors are available to the organism, its inherent variability increases the likelihood that the organism eventually will select and adopt the response which best copes with a given situation.

Fiske (Fiske & Maddi, 1961) asked several relevant questions concerning variability:

1. What is the amount of such variability around the central tendency?
2. Does the amount of variability vary with the value of the central tendency?
3. Does the extent of variability differ for different tests (tasks)?
4. Do individual subjects show different amounts of variability?
5. With what are these individual differences in variability associated?

In an earlier study, the author (Berdie, 1961) found that variability on a task involving advanced high school mathematics appeared consistent within the individual and that this variability might be related to the extent to which a person's college achievement could be predicted on the

basis of aptitude tests. Consequently the present research found its immediate origin in the last of the above listed questions but eventually equally involved questions concerning the consistency and generalizability of intra-individual variability.

The two questions approached here were 1. Can the variability over time of an individual's behavior be reliably and consistently observed? 2. To what extent are persons variable over time on one task also variable on other tasks?

Once these two questions are answered, and if variability over time can be reliably observed and if variability over time is not specific to each task, then the variability of an individual may be a useful concept in understanding his behavior. For example, variable persons may be more or less predictable, may be more or less task oriented, or may be described by other personality characteristics.

Method

Tasks.

The behavior observed here consisted of responses to six of the Repetitive Psychometric Measures (scores on these tests) developed by Moran and Mefford (1959). These authors reported that each of the six tests represented a distinct factor derived from repeated factor analytic studies. The Aiming test (A) consists of fifteen rows each containing twenty circles and the circles are in rows connected sequentially by a line. The subject places the test on a piece of corregated paper and the

task is to punch a hole inside as many circles as possible within ninety seconds without touching the circles. The subjects in this experiment used a stylus consisting of a pencil sized piece of wood with a thin pin point at one end. The test involves the ability to carry out quickly and precisely a series of movements depending on eye-hand coordination and requires careful placement of the stylus point in specified locations. This particular test is not mentioned in French's 1963 report which does refer to each of the other five tests. In a letter from Moran to the author the Aiming test was described as one of several dexterity tests that did not correlate with each other. In their original paper, Moran and Mefferd (1959) quoted French as saying, "The extent to which this specific factor (Aiming) can be generalized to other abilities is unknown." Moran commented that the same form of this test given repeatedly would serve the purpose of a repetitive measure.

The Flexibility of Closure test (FC) requires the subject to copy thirty-six geometric figures into matrices of dots. Each test form contains thirty-six figures and the subject is allowed three minutes. The task, as described by the authors, is to retain the image of a specified configuration despite the influence of other distracting configurations in the perceptual field. The Number Facility test (NF) is similar to French's N factor and consists of ninety problems each requiring the addition of three two digit numbers.

The Perceptual Speed test (PS) requires the identification of well known symbols in a mass of material. This test is not described in French's 1963 volume. The test consists of rows of thirty digits with an encircled digit at the left of each row and the task is to cross out every digit in

the row similar to the encircled digit. The time limit specified by Moran and Mefferd is two and one half minutes but early experience with this test using the subjects employed in this research suggested that too many subjects completed the test with this time limit and the time limit was reduced to one and one half minutes.

The Speed of Closure test (SC) measures the ability to unify an apparently disparate perceptual field into a single percept and it was a marker test in French's 1954 kit but not included in the 1963 kit. Each form consists of twenty-two lines and each line has letters in it apparently arranged at random but containing from two to four four letter words which are to be encircled. The final test, Visualization (V), consists of tangled lines which must be followed visually from their start to their finish. French calls this the "choosing a path test" and lables the factor spatial scanning; he does not call it a visualization factor.

For each of these tests Moran and Mefferd developed twenty different forms with the original intent that the forms would be equivalent. Later study, however, indicated that on each of the tests but Number Facility the alternate forms were reliably different (Moran, Kimble, & Mefferd, 1964) and correction factors were provided for the twenty alternate forms of these five tests. These correction factors were not used in this study, in light of the experimental design.

The test authors reported test-retest reliabilities comparing scores on form one and form two, ranging from .72 to .94. Intercorrelations of the six tests, using only form one, ranged from .09 to .44. Considering

the purposes for which these tests were to be used here, they appeared to be adequately reliable and sufficiently independent from one another.

Sample.

The population from which the subjects were drawn consisted of freshmen entering the University of Minnesota Institute of Technology in the Fall of 1966. All freshmen were informed of the possibility of participating in the experiment and from those who volunteered, subjects were selected on the basis of class schedules, availability of data, and proximity to campus. Each subject was paid a total of forty dollars and subjects appeared well motivated. The subjects tended to be somewhat superior academically to the total Fall quarter entering class because the experiment was conducted during the second quarter and involved only students who survived the Fall quarter and returned. For the most part the subjects consisted of fairly representative bright college students who had survived at least one demanding academic quarter and who were motivated to earn forty dollars by participating in an experiment that would cause them no stress or discomfort.

Procedures.

The experiment was conducted in a well isolated sub-basement room with overhead lights and lamps arranged so that illumination was not brilliant but subjects could see comfortably. No noise from outside the building penetrated the room and little traffic passed in the corridor outside of the door. Temperature in the room was constant and comfortable although when the door was closed there was little ventilation.

Insofar as the subjects remained in the room for periods of only twenty minutes and the door was kept open for at least one half hour between sessions, lack of ventilation produced no discomfort.

Subjects w. e seated in the center of the room in classroom chairs with arm tablets. They were divided into five groups and a group was tested each day at 9:30 a.m., 12:30 noon, 1:30 p.m., 2:30, and 3:30. Assignments to time periods were based on the class schedules submitted by subjects.

Approximately one-fifth of the subjects took form one of the test on the first day, form two on the second, etc. Another group of subjects took form five on the first day, form six on the second day, and on the twentieth day took form four. One group of subjects started on form nine, another group on form thirteen, and another group on form seventeen, in order to provide some randomization of form-sequence influence. Within each time session, students were randomly assigned to sequence groups.

At the first session, the experimenter read to each group an introductory statement and a trained and experienced psychometrist then read the test instructions, administered the practice exercises provided by Moran and Mefferd, and administered the tests.

Testing schedules for each group were arranged Monday through Friday for four successive weeks and subjects who missed sessions made them up during the fifth week. Of the one hundred subjects, sixty-two attended daily and the remaining thirty-eight accounted for a total of seventy-six absences later made up. Twenty-two of these thirty-eight persons were absent once, others were absent from two to eight times. Subjects did not

always attend the session to which they originally had been assigned and sixty-three such time changes were observed out of the two thousand subject-session attendances. Over ninety-five percent of the tests were administered to the subjects at the time of day originally scheduled.

At the last session, at the completion of the last form of the last test, each subject completed a questionnaire reporting his reactions to the tasks and his perceptions of the purpose of the experiment. The subjects were told at the first session that the purpose of the experiment was to compare the psychological characteristics, as measured by these tests, of students in technology and science to those of other students.

Analysis.

The twelve thousand test papers were scored by research assistants and when scoring was completed, two hundred papers for each test were drawn, representing all of the twenty forms, and these were rescored. The original scores were compared to those obtained by rescoring and frequencies of errors of various sizes were tabulated and correlations determined. Scoring was judged to be adequate for five of those six tests but all two thousand of the Number Facility tests were rescored. The test scores were then entered on basic record cards for each student, verified, and then punched and verified on IBM cards.

For each of the six tests a variability index was computed for each student. This consisted of the variance ($S D^2$) of the twenty raw scores derived from the twenty forms of the tests. At the same time, for each of the six tests a mean score was computed for each student, this consisting of the mean of the twenty scores derived from the twenty forms.

Then, in order to facilitate comparisons between tests and to provide a basis for obtaining a total variance index, each raw score was transformed to a standard score, using a mean equivalent to fifty and a standard deviation equivalent to ten, based on the distribution of one hundred scores of each form of each test. For example, the one hundred scores on form one of the Aiming test were selected, the mean and standard deviation calculated, and for each student, his raw score on form one was transformed to a standard score based on this distribution. Then, for each of the six tests a variability index for each student was computed, along with a mean index, based on the twenty standard scores. A seventh variance and mean index were calculated for each student, based on all one hundred and twenty of the scores.

Thus, for each student six variance indices and six mean indices based on raw scores were available, and seven variance indices and seven mean indices were available based on standard scores.

The consistency of the variability index was revealed by what corresponds to an odd/even reliability coefficient. The scores on each of the ten odd numbered forms were used to provide a variance and a mean and the scores on each of the ten even numbered forms provided comparable indices. On each test, each subject had two variance indices and two mean indices and in each instance, the correlation was calculated between the even numbered form and the odd numbered form indices. The analysis was done first using raw scores and then standard scores.

The correlation coefficients then were calculated between the variance indices on the six tests, first using the indices based on raw scores and then the indices based on standard scores. The mean indices were analyzed similarly.

Results

Reliability of Indices.

Table 1 shows the correlations between variance and mean indices based on odd numbered and even numbered forms, using raw scores. Table 2 presents similar information based on standard scores and includes information on the total score, which consists of the sum of the standard scores for the six tests. The test scores are reliable, as shown by the mean score correlations which range from .96 to .99.

Two of the variance indices, one based on Aiming and the other on Number Facility, show relatively high consistency; two, Perceptual Speed and Speed of Closure, provide correlation coefficients in the mid-fifties. Visualization provides the lowest reliability coefficient, .25. The total variance index provides a correlation of .89, suggesting that whatever this is, it is an index that can be obtained rather consistently.

These estimates of reliability are based on ten scores. When we use the Spearman-Brown prophecy formula, reliability estimates of the variability index based on twenty scores are: Aiming, .91; Flexibility of Closure, .58; Number Facility, .89; Perceptual Speed, .71; Speed of Closure, .73; and Visualization, .40. One can conclude that on Aiming and Number Facility variability within persons tends to be remarkably consistent and consistency of variability is found on all other tests.

The author (Berdie, 1961), in a previous study, observed the reliability of a similar variance measure based on ten sub-scores of a mathematics achievement test. The reliability coefficients for various groups

Insert Table 1 about here

Insert Table 2 about here

in that study ranged about .90. Using a somewhat similar method of analyses of varied repeated personality assessment data, Fiske (1957) reported odd/even reliabilities extending from .46 to .96. His results also suggested that the extent of consistency of a person's variability depended in part on the task or instrument used.

Relationships between Variance Indices.

Table 3 shows the intercorrelations between the six variance indices based on raw scores and Table 4 similar intercorrelations of the seven indices, including the total variance index, based on standard scores. In Table 3, of the fifteen correlations, four were significant beyond the .01 level of probability, one between the .05 and .01 level. The variances for Aiming and Number Facility correlated .47, between Aiming and Speed of Closure, .28, between Speed of Closure and Number Facility, .22, and between Number Facility and Visualization, .27. The highest intercorrelation was found between the two variance indices having the highest reliability and the intercorrelations must be examined in light of the reliabilities of the variance indices.

Using the uncorrected reliabilities based on the odd and even numbered forms, and correcting the inter-test variance correlations for attenuation (unreliability), the correlation between the variances for Aiming and

Number Facility increases from .47 to .58, between Aiming and Speed of Closure from .28 to .41, between Number Facility and Speed of Closure from .22 to .33, between Number Facility and Visualization from .27 to .60. These correlations suggest that some of the observed independence between the variances is due to the unreliabilities of the variance indices.

Insert Table 3 about here

Insert Table 4 about here

The reliability coefficients themselves are minimum estimates and one can correct them as we have done before by applying the Spearman-Brown prophecy formula. Unlimited corrections of this sort to statistical data lead to a morass of difficulty, particularly when one is concerned with prediction, but in this instance we are concerned with arriving at some estimate as to relationships between variances, and these should be based on the best reliability estimates. Table 5 shows the intercorrelations of variance indices, where the correlations are corrected for attenuation and the reliability coefficients used have been corrected with the Spearman-Brown prophecy formula. This table gives an optimal estimate of the relationships.

Recognizing the questionable assumptions that have to be made with these two corrections entering into the coefficients, the table reveals that the variance on each of the tests is to some extent related to the variance on one or more of the other tests. The variance index on Number Facility is significantly correlated with the index of each of the other

Insert Table 5 about here

five tests. The variances on Flexibility of Closure and Speed of Closure are related to variances on four of the other five tests. Perceptual Speed variance is correlated to three of the other indices, and two of the coefficients are negative, and Aiming and Visualization variances each are correlated with two of the other variance indices. The Aiming and Number Indices have the highest reliabilities, are the most highly inter-correlated, and the Number Facility index is significantly related with each other index, although the correlations are small. The best indication of the variance domain may be provided by the Number Facility and Aiming tests.

Discussion

These results suggest that intra-individual variability is not specific to each task and neither is there a strongly generalized characteristic of variability that extends over a broad variety of tasks. Rather, the conclusion is that the variability of a person on one task is somewhat related to his variability on certain other tasks and if one is to speak of such variability for a person, one must specify the tasks on which statements are based. If more reliable means can be developed for observing intra-individual variability, better defined clusters of tasks may appear but at present from among the tasks observed here the tasks measured by the Aiming and Number Facility tests provide the best indicators of variability.

The interpretation of these findings depends on other observations and analyses. A series of analyses of variance revealed that the test forms are not equivalent and also that a significant practice or learning effect

was present insofar as on all six of the tests daily mean scores for the group tended to increase from the beginning to the end of the experiment. On five of the tests there was no evidence that the time of day of testing was related to mean scores on the tests, but on the sixth test there was some suggestion that this relationship might exist.

Fiske raised the question regarding the relationship between the variability index and the value of the mean. The correlations between the variance index and mean index for each of the tests here were: A .41, FC .55, NF .30, PS -.26, SC .57, V .03. Five of the coefficients are significant; three are positive and moderately high; one is negative. Examination of two of the bivariate distributions provided no evidence that the variance indices were restricted at the low and high ends of the distributions of mean indices and the relationships appeared rectilinear. The fact that persons with high scores tended to be more variable provides some support for the hypothesis that variability extends the opportunity for the development of adaptive behavior.

The questionnaires completed by the subjects at the end of the experiment suggested that they were well motivated throughout the experiment and eighty-three percent of them reported that they consistently put forth all of their effort in doing as well as they could. Ninety-two percent of them reported that they were able to work on these tests much more effectively on some days than they could on others. Eighty-eight percent reported that on the whole they enjoyed taking the tests. The test they enjoyed least was the Aiming test and the test they enjoyed next least was the Number Facility test, the two that provided the best variance indices.

Immediately after the last form of the last test was administered, the students responded to an open ended question asking what they thought the real purpose of the experiment was. Thirty-three percent provided the explanation given at the beginning of the experiment, that was, to compare technology students to other students. Eighteen percent of the students reported that their perception of the purpose of the experiment was related to the consistency of behavior. Other reported purposes related to describing technology students, learning and improvement, motivation, and eye movements. Only four students reported that they did not know what the purpose of the experiment was. At the end of the questionnaire, students were presented with a check list of five items pertaining to the purpose of the experiment. In responding to this list, fifty-seven percent of the subjects checked the item, "The experiment was concerned with the consistency of my test behavior", thirty-three responded that the purpose was, "To determine how well I did on these tests in relation to my fellow students also taking the tests." These figures suggest that a reasonably large proportion of the subjects had some realization that the experiment was concerned with the consistency of behavior but there was nothing to indicate that the subjects for the most part were strongly motivated to behave consistently.

A supplementary analysis suggested that the indices of intra-individual variability used here are related to the predictability of students' academic behavior. For example, groups divided on the basis of the total variance index derived from the twenty forms of all of the six tests

(120 scores), into high and low variability groups did differ in predictability. The average error of grade point prediction for the high variability group was .28, for the low variability group .009, a difference statistically significant between the .01 and .05 levels.

Conclusions

Large and significant differences are found among individuals in their variability of behavior over time. For example, one subject had a mean score of 43 for the twenty forms of the Number Facility test, with a standard deviation of 1.75, and another subject with the same mean score had a standard deviation of 4.07. On each test large individual differences are found in variability over time.

The reliability with which these differences can be observed varies from task to task and the two tasks providing the most consistent variance index were the Aiming and Number Facility tasks. These were the two tasks that placed the students under the most stress insofar as they were the two least preferred by the subjects.

Variability over time on some tasks is related to variability on other tasks, but these relationships are no more than moderate, even taking into account the relative inadequacies of the means of observation, and the highest correlation between any two of the variance indices was only about .50.

If an easily observable variability characteristic had been identified extending over the six tests, one would face a difficult problem related to the highly speeded nature of the six tests. One then would have to determine the extent to which such a generalized intra-individual variability was related to variability in speed performance, rather than to variability

over different tasks. The relatively small relationships among variances observed here well might be due to the common element of speed characterizing all of the tasks and one might be justified here concluding only that to some extent the speed with which persons perform tasks shows some consistent intra-individual variability, quite apart from the task involved. However, all of the tasks were speeded and if this intra-individual variability were primarily a function of variations in speed, one would expect greater consistency among variance within tasks.

The obtained results suggest that at least two of the tasks studied, Aiming and Number Facility, can provide adequate measures of intra-individual variability. The next question asked earlier by Fiske, is, "With what are these individual differences in variability associated?"

January, 1968

References

Berdie, R. F. Intra-individual variability and predictability. Educ. psychol. Measmt., 1961, 21, 663-676.

Fiske, D. W. An intensive study of variability scores. Educ. psychol. Measmt., 1957, 17, 453-465.

Fiske, D. W. & Maddi, S. R. Functions of varied experiences. Homewood, Ill.: The Dorsey Press, Inc., 1961.

French, J. W. (Ed.) Manual for kit of selected tests for reference aptitude and achievement factors. Princeton, N. J.: Educational Testing Service, 1954, Revised 1963.

Hull, C. L. Variability in amount of different traits possessed by the individual. J. educ. Psychol., 1927, 18, 97-104.

Moran, L. J., Kimble, J. P. Jr. & Mefferd, R. B. Jr. Repetitive psychometric measures: equating. Psychol. Rep., 1964, 14, 335-338.

Moran, L. J. & Mefferd, R. B. Jr. Repetitive psychometric measures. Psychol. Rep., 1959, 5, 269-275.

Wechsler, D. The range of human capacities. (2nd Ed.) Baltimore: Williams and Wilkins, 1952.

Table 1

Correlations for 100 University of Minnesota Institute of Technology Freshmen on Variance Index and Mean Raw Score Based on Ten Odd Numbered Forms and Ten Even Numbered Forms of Each of Six RPM Tests

	Variance Index					Mean Score				
		Odd Number		Even Number			Odd Number		Even Number	
	r	M	SD	M	SD	r	M	SD	M	SD
Aiming (A)	.83	210.74	182.04	199.94	186.46	.96	111.40	15.64	114.34	16.13
Flexibility of Closure (FC)	.41	20.16	11.79	17.55	10.57	.97	18.29	4.77	20.10	5.16
Number Facility (NF)	.80	29.54	29.45	24.44	26.11	.99	43.04	10.34	44.46	10.59
Perceptual Speed (PS)	.55	33.28	20.63	44.12	30.67	.97	61.64	5.53	63.84	6.18
Speed of Closure (SC)	.57	69.36	33.27	52.85	28.14	.98	38.85	6.43	39.31	6.38
Visualization (V)	.25	47.67	25.27	38.31	23.55	.98	53.84	7.80	54.28	8.02

All correlations significant beyond .01 level except onesignificant between .01 and .05.

Table 2

Correlations for 100 University of Minnesota Institute of Technology Freshmen of Variance Index and Mean T Score Based on Ten Odd Numbered and Ten Even Numbered Forms for Six RPM Tests and for "Total Measure"

	Variance Index					Mean Score				
		Odd Number		Even Number			Odd Number		Even Number	
	r	M	SD	M	SD	r	M	SD	M	SD
Aiming (A)	.82	47.60	40.04	44.65	41.12	.96	49.99	7.53	50.00	7.71
Flexibility of Closure (FC)	.42	47.03	27.46	37.28	19.04	.97	50.00	7.56	50.00	8.13
Number Facility (NF)	.80	20.68	21.59	17.05	19.19	.99	50.00	9.01	50.00	9.19
Perceptual Speed (P.	.53	42.84	33.91	37.42	32.75	.97	50.00	7.81	50.00	8.12
Speed of Closure (SC)	.71	61.42	30.35	52.33	27.04	.98	50.00	6.65	50.00	7.24
Visualization (V)	.37	39.26	22.26	34.38	21.20	.98	50.00	8.02	50.00	8.29
"Total Measure"	.89	66.91	26.56	63.15	27.37	.99	50.00	5.79	50.00	6.10

All correlations significant beyond .01 level.

Table 3

Intercorrelations Between the Six "Variance" Indices (Using Raw Scores) for 100 University of Minnesota Institute of Technology Freshmen Each Taking Twenty Forms of Each of Six RPM Tests.

	FC	NF	PS	SC	V	Mean	SD
Aiming (A)	-.02	.47**	.01	.28**	.01	202.04	167.48
Flexibility of Closure (FC)		.17	-.15	.16	.32**	19.09	9.45
Number Facility (NF)			-.18	.22*	.27**	26.58	25.11
Perceptual Speed (PS)				.14	-.01	38.11	21.34
Speed of Closure (SC)					.06	58.45	25.88
Visualization (V)						41.59	18.23

* P <.05

** P <.01

Table 4

Intercorrelations Between the Seven "Variance" Indices (Using T Scores) for 100 University of Minnesota Institute of Technology Freshmen Each Taking Twenty Forms of Each of Six RPM Tests (Also Included is the "Total Variance" Index).

	FC	NF	PS	SC	V	TV	Mean	SD
Aiming (A)	.00	.49**	-.06	.26**	.05	.42**	44.89	36.62
Flexibility of Closure (FC)		.14	-.10	.22*	.27**	.15	40.84	19.25
Number Facility (NF)			-.18	.20*	.32**	.50**	18.23	18.38
Perceptual Speed (PS)				-.01	.03	-.03	39.15	28.02
Speed of Closure (SC)					.03	.21*	54.60	25.26
Visualization (V)						.28**	35.76	16.97
Total Variance (TV)							64.69	25.97

* P <.05

** P <.01

Table 5

Intercorrelations Between the Six Variance Indices Using Raw Scores, Correlations Corrected for Attenuation (Unreliability of Variance Indices) Using Spearman-Brown Corrected Reliability Coefficients.

(N=100)

	FC	NF	PS	SC	V
Aiming (A)	-.03	.53	.01	.35	.02
Flexibility of Closure (FC)		.24	-.23	.25	.67
Number Facility (NF)			-.23	.28	.46
Perceptual Speed (PS)				.20	-.02
Speed of Closure (SC)					.11

Intra-individual Temporal Variability

and Predictability(1)

Ralph F. Berdie

Office of the Dean of Students

University of Minnesota

Although the assumption that behavior is predicatable does not require that behavior be consistent, consistency of behavior may facilitate its predictability. If some person's behaviors are more consistent than those of other persons, they may be more predictable. If one person leaves his apartment for work every morning exactly at eight o'clock, and another person's time of departure varies widely between seven-thirty and eight-thirty, then predictions regarding when the men would return from lunch, based on the time they left for lunch, presumedly might be more accurate for the former than the latter individual. Persons whose habits are regular, or whose behavior is consistent over time, may be the most predictable. An informative review of the research on predictability has been presented by Tolbert (1966). Fiske and Maddi (1961) discuss in detail the concepts of variability.

[1]This study was supported with funds from the U.S.O.E.-O.E.C.-3-7-068694-2082. Appreciation is expressed to Mr. Richard Arvey and Miss Diane Johnson for their contribution in analyzing the data.

Only suggestive evidence pertains to the relationships between consistency and predictability and two studies provide some insight. This author (Berdie, 1961) reported that the correlations between college aptitude test score and first quarter college grade point average were .70 for a group of students with little variability shown on responses to a mathematics test and .44 for a matched group with high variability. Most of the analyses in that study provided results that failed to attain statistical significance, but the results were consistent and the conclusion was that intra-individual differences in variability were perhaps related to predictability. A later duplication of this study with somewhat similar methods provided no confirmation.

A more recent unpublished study by Arvey showed that 200 college students could be divided into two groups on the basis of the variability of their high school grades. The correlation between predicted college grade point average based on test scores and high school rank, and obtained grade point average for the most variable group was .10, for the least variable group .69. Similar correlations based on an independent cross-validation sample were .23 and .74. Individuals whose behavior was more variable in high school were less predictable in college and intra-individual variability was regarded as a practical moderator variable. (Saunders, 1956)

The use of a moderator variable in prediction depends on the purposes of prediction and the situation. Sex often is a moderator variable and academic prediction for girls tend to be more accurate than for boys (Seashore, 1962). A multiple correlation between first year grades and three predictors was .46 for men and .57 for women in the University of Minnesota Arts College. For the total group, $R = .52$. Although the moderator variable

could provide better predictive efficiency for one group, it really did not contribute to selection for the total group. For predictive purposes in advising and counseling, however, it did effect probability statements, and it did provide incentive for further analysis of the prediction process.

If the classification of persons on the basis of their consistency of performance is to aid in improving predictions for sub-groups of these persons, then the question of the generalizability of such consistency becomes important. To what extent can consistency of behavior be regarded as a trait which characterizes the individual?

Using seven intra-individual variability indices based on six tests, the author (Berdie, 1967) found that the intercorrelations between the indices ranged from -.18 to .49. Of the fifteen correlations between variance indices of the six tests, five attained statistical significance beyond the .05 level and the variance indices for five of the tests were significantly correlated with the variance index of at least one of the other tests. For one test the variance index was not correlated with the index of any of the other tests. On these six repeated psychological measurements, the evidence suggested that consistency of behavior is not highly specific to each task, but also that currently one cannot conclude that a highly generalized consistency of behavior is observable over several tasks.

Arvey observed that the correlations between variance estimates of high school grades in six subject matter areas, including a total variance, ranged from .04 to .34. Of the fifteen intercorrelations, including the part-whole correlations, nine were statistically significant behond the .05 level. These results suggested that persons whose grades are variable in

high school in one subject tend to obtain variable grades in other subjects but the relationships are low and reflect no easily observed generalized variability trait.

Before one can study the usefulness of indices of intra-individual variability as moderator variables to improve prediction, the behavior or behaviors on which the variability indices are to be based must be selected. The present research was concerned first, with the identification of appropriate variability indices, and secondly, with the relationships between these indices and predictability of academic behavior. The first concern is discussed elsewhere (Berdie, 1967). The purpose of this report is to describe the methods explored in studying relationships between variability and predictability and present the evidence revealed.

Method

Subjects

The subjects were Institute of Technology freshmen entering the University of Minnesota in the Fall of 1966. All freshmen were told of the possibility of participating in the experiment and subjects were selected from volunteers on the basis of class schedule, proximity to campus, and availability of data. The subjects tended to be somewhat superior academically to the total entering class because the experiment was conducted during the second academic quarter and included only students who survived the first quarter. These students were a representative group of bright college students motivated to earn forty dollars by participating in an experiment that would cause them no stress or discomfort.

Tasks Providing a Basis for Inferring Intra-indvidual Variability

Subjects were classified in terms of intra-individual variability on the basis of their performance on six of the brief, highly speeded, Repetitive Psychometric Measures developed by Moran and Mefferd (1959). These are relatively pure factorial tasks frequently described in the factor analytic literature.

The Aiming test (A) consists of fifteen rows each containing twenty circles connected sequentially by a line. The subject places the test on a piece of corregated paper and punches holes inside as many circles as possible without touching the circle. Subjects used a stylus consisting of a pencil sized piece of wood with a thin metal point. The Flexibility of Closure test (FC) requires the subject to copy thirty-six geometric figures into matrices of dots. The task, as described by the authors, is to retain the image of a specified configuration despite the influence of other distracting configurations in the perceptual field. The Number Facility test (NF) consists of ninety problems each requiring the addition of three two digit numbers.

The Perceptual Speed test (PS) consists of rows of thirty digits with a digit in the left hand column of each row encircled and the task is to cross out every digit in the row similar to the encircled digit. The time limit specified by Moran and Mefferd for this test is two and one half minutes but early experience suggested that too many subjects completed the test with this time limit and the time limit was reduced to one and one half minutes.

The Speed of Closure test (SC) measures the ability to unify an apparently disparate perceptual field into a single percept. Each form consists of twenty-two lines and each one has letters in it apparently arranged at random but containing from two to four four-letter words which are to be encircled. The final test, Visualization (V) consists of tangled lines which must be followed visually from their start to their finish.

For each of these tests Moran and Mefferd developed twenty different forms so that the forms would be equivalent. Later research indicated that for five of the six tests the alternate forms were reliably different (Moran, Kimble, and Mefferd, 1964) and correction factors were provided for the twenty alternate forms of the five tests. These correction factors were not used in this study in light of the experimental design employed.

Moran and Mefferd reported reliability coefficients, based on correlations between form one and form two, for the six tests ranging from .72 to .94. For each of the one hundred subjects included in the present research a score was derived from the odd numbered forms of each test and from the even numbered forms of each test and the correlations between these scores ranged from .96 to .99. The tests appeared to be reasonably reliable.

Moran and Mefferd reported that intercorrelations between the six tests, using only one form, ranged from .09 to .44. In the present experiment a total score was available for each subject on each test, based on the twenty forms of the test, and the intercorrelations between these scores ranged from .28 to .65, considerably higher than the intercorrelations reported by the original authors. Their correlations were based on scores

derived from tests requiring from two to five minutes and involved measurements of considerable less reliability than characterizing the scores used here. This raises an interesting question as to the extent to which intercorrelations found between tests and resulting factor analytic results depend on the length and related reliabilities of the tests included in the factor analyses. In spite of the higher intercorrelations found here, the tests appear to be sufficiently independent from one another to justify their use in this research.

Procedures

The one hundred subjects were divided into five groups and each group was tested at the same time each day for five days during four successive weeks. One group was tested at 9:30 a.m., another group at 12:30 noon, and the three remaining groups tested at later times during the afternoon. Assignments to time periods were based on students' class schedules.

About one fifth of the subjects took form one of the test on the first day, form two on the second, etc. Another group of subjects took form five on the first day, form six on the second day, and on the twentieth day took form four. One group of subjects started on form nine, another group on form thirteen, and the final group on form seventeen. Within each time session students were randomly assigned to sequence groups.

The experimenter read to each group an introductory statement to explain the purposes of the research. Subjects were told the experiment was designed to provide comparisons between the psychological characteristics as measured by these tests of students in technology and science with

characteristics of other students. The tests were administered by a trained and experienced psychometrist who read the test instructions at the first session, administered the practice exercises provided by Moran and Mefferd, and then administered the tests. Testing schedules for each group were arranged Monday through Friday for four successive weeks and subjects who missed sessions made them up during the fifth week. Over ninety-five percent of the tests were administered to the subjects at the time of day originally scheduled and of the two thousand subject-attendances, 1,924 occurred on the day scheduled.

At the completion of the last form of the last test, each subject completed a questionnaire reporting his reaction to the tasks and his perceptions of the purpose of the experiment. In response to an open-ended question, 18 percent of the students reported that they thought that the purpose of the experiment was related to the consistency of behavior but later, in responding to a check list containing five items pertaining to the research purpose, 57 percent checked the item, "The experiment was concerned with the consistency of my test behavior." These figures suggested that a reasonably large proportion of subjects had some realization that the experiment was concerned with the consistency of behavior but little indicated that the subjects were strongly motivated to behave consistently. They did report high motivation to perform well.

The experiment was conducted in a well isolated sub-basement room with over-head lights and lamps arranged so that illumination was adequate. Little distraction occurred. Subjects were seated in the center of the room in classroom chairs with arm tablets.

Prediction Data

In addition to the one hundred and twenty Repetitive Psychological Measurement test scores, additional data were available for subjects. Those included a score on the Minnesota Mathematics test, a score on the Minnesota Scholastic Aptitude test, high school percentile rank, obtained Fall quarter grade point average, and predicted Fall quarter grade point average derived from the mathematics test. From the last two indices a discrepancy score was obtained by subtracting the obtained grade point average from the predicted grade point average.

The Minnesota Mathematics test was developed as a means for admitting students to the University of Minnesota Institute of Technology and is a comprehensive examination covering high school mathematics, with emphasis on algebra. Correlations between this test and Fall quarter grade point average range about .50 for groups of freshmen admitted to college on the basis of information other than this test score. The Fall quarter grade point average was based on the grades of all courses taken during the first quarter in the Institute of Technology. Students customarily register for from three to four courses requiring a total of fifteen to twenty hours per week of class and laboratory attendance. For the one hundred subjects the correlation between the Minnesota Mathematics test score and the obtained Fall quarter grade point average was .33 and the results of the experiment must be interpreted in light of the relatively low predictive efficiency of this instrument.

For each student the predicted Fall quarter grade point average was obtained by using a single variable regression equation based on the mathematics test scores derived from a previous class of Institute of Technology freshmen. For the one hundred subjects the correlation between the mathematics test score and the predicted Fall quarter grade point average was .84. The difference between this and a correlation of 1 reflects possible change in the population and the extent to which error has been incorporated into the original regression equation. The reliability of the obtained Fall quarter grade point average can be inferred from the correlation between obtained Fall quarter grade point average and obtained Winter quarter grade point average, .70.

Analysis

The twelve thousand Repeated Psychological Measurement tests were scored by research assistants; scoring was checked; and when necessary papers were re-scored. Test scores and other data then were entered on basic record cards for each student, verified, and then punched and verified on IBM cards.

For each of the six tests a variability index was computed for each student. This consisted of the variance (SD^2) of the twenty raw scores derived from the twenty forms of the tests. Then, in order to facilitate comparisons between tests and to provide a basis for obtaining a total variance index, each raw score was transformed to a standard score, using a mean equivalent to 50 and a standard deviation equivalent to 10, based

on the distribution of one hundred scores of each form of each test. To illustrate this, the mean and standard deviation were calculated for the one hundred scores on form one of the Aiming test and for each student, the raw score on form one was transformed to a standard score based on this distribution. Using the standard scores for each of the six tests variability indices were computed and a seventh variance index was calculated for each student, based on all one hundred and twenty standard scores.

The consistency of the variability indices was analyzed and an odd/even reliability coefficient obtained for each of the six tests and for the total variance index. The raw scores on each of the ten odd-numbered forms were used to obtain one variance index and the scores on each of the ten even-numbered forms were used to obtain a second index and the correlations obtained for the six tests were: A .83, FC .41, NF .80, PS .55, SC .57, V .25. The comparable correlation for the variance index encompassing all 120 standard scores was .89.

The variance indices based on five of these tests and on the total score were sufficiently reliable to suggest that these measures of variability themselves were consistent.

Table 1 presents the intercorrelations between the six variance indices. Six of the fifteen correlations are statistically significant and the variance on five of the tests is related to the variance on at least one other test. The variances based on the six tests and the total variance, which provides a part-whole correlation, correlate between -.03 and .49.

These intercorrelations suggest that variability over time on some tasks is related to variability on other tasks, but these relationships are no more than moderate, even when one considers the reliabilities of the observations. Intra-individual variability is not specific to each task

and neither is there a strongly generalized characteristic of variability that extends over a broad variety of tasks.

These results suggest that at least two of the tasks studied, Aiming and Number Facility, and the total variance index, may provide adequate measures of intra-individual variability.

The most comprehensive variance index, and the one with the highest reliability, was the one based on all one hundred and twenty standard scores. Although this measure of intra-individual variability encompassing six different tasks is difficult to interpret because of its heterogeneity, it does provide one means for testing the hypothesis that intra-individual variability is related to predictability of behavior.

To test this hypothesis, the mean of each student's one hundred and twenty standard scores and the variance of these scores were recorded and the one hundred subjects were divided into four groups, those with both the means and the variances below and above the group average, those with the means below and the variances above the group average, and those with the means above and the variances below the average. Then groups were recombined so that one group included all of those with mean indices above the group average, one with mean indices below the group average, one with variance indices above the group average, and one with variance indices below the group average.

For each of these eight groups statistics were calculated, including: the means for the mathematics test, the predicted Fall quarter grade point average, the obtained Fall quarter grade point average, the discrepancy

between these, the correlations between the mathematics test and obtained Fall quarter grade point average and predicted grade point average and obtained grade point average.

Then for the total group of one hundred subjects correlations were calculated between all of the variables, including the correlations between the several variance indices and discrepancy between obtained and predicted grade point average.

Then, in light of the demonstrated interactions between the variance index and the mean index, the total group of one hundred was divided into four sub-groups: students with correctly predicted grade point averages, which were high, those with correctly predicted grade point averages which were low, those whose grade point averages were under predicted, and those whose grade point averages were over predicted. The variance indices on the six tests were compared for these groups (Hobert & Dunnette, 1967).

Then the group was divided into four other sub-groups consisting of subjects with both low mean scores and low variance indices on the Aiming test, those with low mean scores and high variance indices, those with high mean scores and low variance indices, and those with high mean scores and high variance indices, and the grade point discrepancies for these groups were compared.

Then the entire group was divided into two groups, one with above average variance indices on the Number Facility test and one with below average variance indices on this test, and the groups were compared on the basis of correlations and means.

In order to allow for possible effect of size of raw scores on the tests, another group of students was selected consisting of those who had mean scores on both the Aiming and Number Facility tests that placed them within plus and minus one standard deviation of the means for the total groups and for this group correlations were determined between variance indices and other characteristics, including grade point discrepancies.

Next, for each student the coefficient of variation was calculated on the Number Facility test and comparisons were made of students with coefficients of variation above the median and those below the median.

These various analyses were designed to provide information concerning the differences in predictability of students with different variability characteristics.

Results

Consideration of Variance and Mean Indices

Table 2 shows the characteristics of the one hundred subjects divided into eight sub-groups on the basis of individual subject mean and variance on total score derived from responses to all twenty forms of each of the six tests. The first group of twenty-seven subjects had a mean standard score for the one hundred and twenty tests and a variance of these scores placing them at or above the group mean average and group mean variance for the total group of one hundred subjects. The second group had mean scores above the group average but variance indices below the group average and the next two groups consist of thirty students with both means and variances below the group average and twenty-two with means below and variances above the group average. The fifth, the high mean group, consists

of the first and second groups combined, the next group, the low mean group, of the third and fourth groups combined, the next "high variance" group of the first and fourth groups combined, and the last "low variance" group of the second and third groups combined.

The hypothesis, "Subjects who were more variable are less predictable" can be tested by looking at the correlations between predictors and criteria and also by looking at the discrepancies between predicted and obtained grade point averages. For the low variability group, the eighth group, the correlation between mathematics test score and the obtained Fall quarter grade point average was .39, as compared to the correlation of .28 for the high variability group, group seven. The difference is in the expected direction but is not statistically significant.

The mean discrepancy between predicted and obtained grade point average for the low variability group was .01 with a standard deviation of .63, as compared to the mean discrepancy of .28, with a standard deviation of .61, for the high variability group. The "t" for the mean difference was 2.19, showing that the difference was statistically significant between the probability levels of .05 and .01. The discrepancy analysis supports the hypothesis that persons who have low variability indices will perform more as predicted compared with persons with high variability.

The two variability groups were quite similar on the basis of mean mathematics test scores and predicted grade point average. The high variability group obtained somewhat better grades on the average than did the low variability group but the difference was not statistically significant.

The correlational comparisons of the four groups divided on the basis of both variance and means reveal for the groups with above average mean indices, that the low variability group is more predictable than the high variability group. The correlations between the mathematics test and the Fall quarter obtained grade point average for the former group was .67, for the latter group .16. Transformation of these coefficients to "z" and evaluation of the significance of the difference provide a "t" of 2.08, significant with a probability between .05 and .01. When the mean discrepancies between predicted and obtained grade point averages are compared, .14 and .32, the differences are not statistically significant.

When the low and high variability groups with below average mean indices are compared, the correlations are almost identical, .35 and .36, and the discrepancies .01 and .23, are not statistically significantly different. Thus, the hypothesis has some support from this analysis but the results suggest that the hypothesized relationship between variability and predictability is observable mainly for subjects achieving high mean scores on the tasks from which the variability index is derived.

Intercorrelations of all Variables

Table 1 shows the means, standard deviations, and intercorrelations for all the variables, with the RPM indices based on standard scores. A similar table using raw scores revealed essentially the same relationships. Obtained Fall quarter grade point average, the criterion used here, was substantially correlated with discrepancy between obtained and predicted grade point average and significantly correlated with mean scores on the

Aiming, Number Facility, Perceptual Speed, Visualization, and total mean score on the RPM tests. Obtained Fall quarter grade point average was not correlated significantly with any of the variance indices.

The GPA discrepancy was correlated .19, with the variance index on the Visualization test, but this was not statistically significant. The discrepancy was correlated with the mean score on three of the six tests and with the total mean score. The absence of significant correlation between the discrepancy and the seven variance indices does not lend support to the general hypothesis and in light of the results presented in Table 2, again suggests that the relationships, if present, are quite complex.

The mean indices are significantly correlated with the variance indices on five of the six tests and this correlation for the Perceptual Speed test is moderate and negative. For the total index the correlation between variance and mean is .21, barely significant at the .05 level.

Consideration of Relationship Between Predictor and Criterion

The relationship between predicted and obtained grade point average and the seven mean and seven variance indices is analyzed in another way in Table 3. A scatter diagram was prepared showing the relationship between predicted and obtained grade point averages and the group mean predicted and group mean obtained averages were determined. The total group was divided into four sub-groups corresponding to the four quadrants. One quadrant contained thirty-one subjects with both predicted and obtained high grade point averages, another contained thirty-six subjects with both

predicted and obtained low grade point averages. One quadrant contained eighteen subjects whose obtained predicted grade point average was high and predicted grade point average low, and the forth quadrant contained the fifteen subjects with high predicted grade point averages and low obtained grade point averages. For each of these groups, Table 3 presents the mean and standard deviation for each of the seven variance indices and for each of the seven mean indices. The subjects with the most accurately made predictions are in the first two groups, the ones with the least accurate predictions in the last two groups. Examination of the average variance indices shows few consistant or meaningful trends. When one combines the sixty-seven subjects in the two groups least accurately predicted, the correct prediction group has the highest mean variance on the Aiming, Flexibility of Closure, and Number Facility tests and the incorrect prediction group has the highest mean variance on the Perceptual Speed, Speed of Closure, and Visualization tests. The mean variances for the two groups on the total measure are practically identical.

On the first three tests the mean variances for the low achievers who were correctly predicted are higher than those of the other three groups, but this is not found for the last three tests. When the two groups with low predicted grade point averages are compared, the group that achieved according to prediction and the group that did better than predicted, a difference in the mean variance on the Aiming test is statistically significant with a probability of less than .01 and the variances also are significantly different for the variance index on this test but

this does not appear to fit in with any of the other results. Perhaps if any differences can be observed in this table, they refer more to differences between high and low achievers than to differences between correct and incorrect predictions.

Consideration of Mean and Variance Indices on Aiming

Table 4 presents the characteristics of four groups divided on the basis of variance index and mean index on the Aiming test, using raw scores. Thirty subjects had both mean and variance indices below the group mean; eighteen had mean indices below and variance indices above the group means, etc. When the two groups of students with low mean indices are compared, those with low variance and those with high variance indices, the difference in the mean discrepancies is statistically significant beyond the .05 level with the high variance group having the greatest discrepancy. When the two groups with high mean indices are compared, again the high variance group has the greatest discrepancy, but the difference is not statistically significant. In Table 2, the statistically significant difference between the high and low variability groups was found for the groups with mean indices above the group average; with the Aiming test, the significant difference was found for the low mean index group. Again some, but far from conclusive evidence is available concerning the relationship between variance and predictability.

Variability on Number Facility

Table 5 shows characteristics of the subjects divided on the basis of variance index on the Number Facility test. Of the ninety-three subjects

for whom complete data were available for this analysis, forty-seven were in the high variability group and forty-six in the low variability group. For these two groups correlations between the predictor test and obtained grade point average were almost identical. The discrepancy between obtained and predicted grade point average was slightly greater for the variable group than for the less variable group, but the difference was not significant.

Middle Range Subjects

Table 6 shows the characteristics of students who on both the Aiming and Number Facility tests had mean scores that placed them within plus and minus one standard deviation for the total group mean. Again no significant correlation appeared between the predicted and obtained grade point average and the variance indices of these two tests. When the twenty-eight subjects with lowest variances on the Aiming test were compared with the twenty-eight subjects with highest variance, the low variability group had the highest correlations between predicted and obtained grade point average, a correlation of .58 as compared to a correlation of .40 for the other group. When the groups were divided on the basis of variability on the Number Facility test, again the low variance group had the highest correlation, .58, as compared to the correlation of .46 for the other group. With groups as small as these, differences between correlations to be significant would have to be much larger than one could reasonably expect to obtain so the results are only mildly suggestive.

Relative Variability

Finally, in order to study the possible effect of the relationships between the mean and variance indices themselves, the use of the coefficient

of variation was explored and the results are presented in Table 7.

The ninety-three students for whom complete data were available were divided on the basis of high and low coefficients of variation on the Number Facility test. For the variable group the correlation between predicted and obtained grade point average was .26, for the less variable group .55. The correlation between the Minnesota Scholastic Aptitude test and grade point average for the variable group was .10, for the less variable group .35. The correlation between the Minnesota Scholastic Aptitude test and high school percentile rank, also differed although not significantly. The only correlations that differed significantly were those between the Mathematics test and high school percentile rank, where the correlation was -.05 for the variable group, +.34 for the less variable group. Using this variability index the two groups also differed on the basis of mean Fall quarter grade point average, with the less variable group achieving better grades.

Discussion

Answering the question, "Is intra-individual variability related to predictability?" is not easy. The definition and measurement of intra-individual variability are complex. The answer to the question depends in part on the tasks used and the derived indices. The function being predicted makes a difference, and more than one method is available for observing the accuracy of prediction. On five of the six tasks used here, the subject's level of performance was significantly related to his variance

and one of these coefficients was negative. The answer to the question also depends in part on the segment of the range being observed of the performance level of the task providing the variance index. Some of the evidence suggests that the answer depends in part on the level being considered of both the predictor variable and the predicted variable. The results do suggest that predictability of behavior, as defined here, is influenced to some extent by the subject's intra-individual variability but that the relationship is not a straight-forward one.

The total variance index based on one hundred and twenty scores is most difficult to interpret insofar as it consists of twenty scores on each of six relatively independent tests and these are all transformed to standard scores. Using this complex index of intra-individual variability, the correlation for the less variable group between the predictor and the criterion is higher, but not statistically significantly so, than the correlation for the variable group. When the mean discrepancies between predicted and obtained grade point averages are compared for the high and low variability groups, however, the difference is statistically significant. In light of this statistically significant difference and in light of the difference in the correlations being in the expected direction, the evidence does suggest that this complex variance index is related somewhat to predictability.

Whereas the correlation between the predictor and predicted variable appears to be relatively easy to interpret, the GPA discrepancy measure

is not. Two methods are available for handling the discrepancy: it can be treated algebraically or arithmetically. Treating it algebraically pays attention not only to the size of the discrepancy but also to its direction whereas treating it arithmetically reflects only the size of the discrepancy and ignore its direction.

The anlysis in Table 2 was done using both the algebraic and the absolute discrepancies. Whereas the difference between the two groups on the algebraic discrepancy was significant, the absolute discrepancies were the same for the groups. When the two groups with the high means were considered, the one with the high variance and the other with the low variance, the algebraic discrepancy was higher for the high variance group. For the two groups with low mean indices, the low variance group had the lowest algebraic discrepancy but the highest absolute discrepancy.

If relationships between the algebraic discrepancy and the other variables were curvilinear a more precise picture of the relationship might be provided by using the absolute or arithmetic discrepancy. These relationships were examined for curvilinearity, and no evidence was observed suggesting other than rectilinear relationships. At the same time these observations were made, the relationships between mean and variance indices were observed and no evidence was found that these were curvilinear.

The difference in correlations shown in Table 2 for the two groups with high means, as compared to the corresponding difference for the two groups with low means, is difficult to explain. For the two groups with

high means, the group with high variance had a standard deviation of 8.20 on the Mathematics test and .52 on the criterion Grade Point Average. The low variance group had a much smaller standard deviation on the Mathematics test, 4.72, and about the same standard deviation on the GPA, .54. The interaction between the mean value and the variance value apparently has to be considered.

If the correlations between GPA discrepancy and the seven variance indices in Table 1 had proved to be statistically significant, the most easily interpreted evidence concerning the hypothesis would be available. Obviously, the relationships were not that simple.

The attempt to consider the interaction between the mean index and the variance index by using the coefficient of variation provided results that indicated, although often not significantly, that taking into account both of the measures, the less variable group was more predictable. The coefficient of variability is not a rigorous statistic but within the defined limits of this analysis, the results did suggest that the hypothesis might well be best tested taking into account measures corresponding both to the mean and variance indices.

Conclusion

The evidence suggests a relationship between the variability over time of a person's behavior on simple tasks and the effectiveness with which his academic performance can be predicted. The relationship between intra-individual variability and predictability is complex and the measurement of

this relationship must consider the subject's level of performance on the tasks from which the variability indices are derived, the level of performance on the predictor and predicted variable, and the method used in determining accuracy of prediction.

If intra-individual variability is an effective moderator variable, accuracy of prediction perhaps can be increased for the least variable group, but prediction for the variable group will remain inefficient. If the efficiency of prediction is to be increased for the total group, and particularly for the group with greater intra-individual variability, then new prediction models may have to be devised and new predictors or new criteria studied.

For example, for the group of persons with high variability, highly speeded tests may be inappropriate as predictors. For this group a rectilinear regression model may be inappropriate and the nature of the shape of the relationship between predictors and criterion may be different. For the least variable group, a criterion based on one quarter of academic performance may be appropriate whereas for the group with greater temporal variability, a criterion of academic success based on several semesters of work may be more appropriate.

Temporal intra-individual variability may prove to be a useful and meaningful concept but before this can be determined further, more efficient means will be needed for observing such variability. The results here can be interpreted only as being suggestive. They are promising.

February 5, 1968

References

Berdie, R. F. The consistency and generalizability of intra-individual variability. in press, 1967.

Berdie, R. F. Intra-individual variability and predictability. Educ. psychol. Measmt., 1961, 21, 663-676.

Fiske, D. W. & Maddi, S. R. Functions of Varied Experience. Homewood, Ill.: The Dorsey Press, Inc. 1961.

Hobert, R., & Dunnette, M. D. Development of moderator variables to enhance the prediction of managerial effectiveness. J. appl. Psychol., 1967, 51, 50-64.

Moran, L. J., Kimble, J. P., Jr., & Mefferd, R. B., Jr. Repetitive psychometric measures: equating. Psychol. Rep., 1964, 14, 335-338.

Moran, L. J., & Mefferd, R. B., Jr. Repetitive psychometric measures. Psychol. Rep., 1959, 5, 269-275.

Saunders, D. R. Moderator variables in prediction. Educ. psychol. Measmt., 1956, 16, 209-222.

Seashore, H. G. Women are more predictable than men. J. counsel. Psychol, 1962, 9, 261-270.

Tolbert, E. L. Predictability in testing and counseling. Personnel & guidance J., 1966, 45, 219-226.

Table 1

Intercorrelations Between Variables Using RPM Standard Scores

N=100

Variable

Variable	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Mean	S D
1. Minn. Math Score	.84	.33	.27	-.11	-.21	-.02	-.11	-.01	.10	.03	-.10	-.26	.14	.09	-.04	-.03	-.02	-.02	32.93	7.86
2. Predicted fall qt. GPA		.39	.26	-.14	-.16	.01	-.10	-.02	.11	-.03	-.10	-.13	.22	.12	-.05	.03	.04	.08	2.21	.35
3. Obtained fall qt. GPA			.70	.86	-.11	.07	-.03	-.03	.14	.16	.04	.28	.14	.23	.26	.12	.26	.29	2.39	.67
4. Obtained winter qt. GPA				.60	-.02	.00	-.05	-.05	.09	-.03	-.01	.19	-.02	.06	.18	.05	.16	.14	2.35	.80
5. GPA discrepancy					-.03	.07	.02	-.02	.09	.19	.10	.38	.03	.18	.26	.11	.26	.27	..18	.62
6. Aiming-variance						.00	.49	-.06	.26	.05	.4	40	.00	-.02	.16	.22	.12	.18	44.89	36.61
7. Flex. Clos. variance							.14	-.10	.22	.27	.15	.31	.55	.20	.33	.29	.38	.45	40.84	19.25
8. No. Facility variance								-.18	.20	.32	.49	.36	.07	.27	.24	.32	.19	.32	18.23	18.38
9. Perc. speed variance									-.01	.03	-.03	-.21	-.26	-.25	-.41	-.33	-.34	-.40	39.15	28.02
10. Speed closure variance										.03	.20	.36	.25	.17	.38	.55	.32	.44	54.60	25.26
11. Visualization variance											.28	.16	.09	.26	.09	.16	.05	.18	35.75	16.97

continued

Table 1 (continued)

Variable

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	Mean	S D
12. Total variance												.26	.06	.16	.08	.19	.20	.21	64.69	25.97
13. Aiming-mean													.39	.37	.55	.53	.59	.75	49.99	7.54
14. Flex. Clos. mean														.26	.39	.43	.59	.68	49.99	7.79
15. No. Facility mean															.42	.36	.50	.67	49.99	9.08
16. Perc. speed mean																.64	.66	.80	49.99	7.90
17. Speed closure mean																	.55	.76	49.99	6.90
18. Visualization																		.86	49.99	8.10
19. Total mean																			49.99	5.93

r of .20 significant at .05 level

r of .26 significant at .01 level

Table 2

Characteristics of Subjects Divided into Groups on the Basis of Means on the Standard Score

Total Mean Index and Total Variance Index.

Mean index	hi	hi	lo	lo	hi	lo			Total
Variance index	hi	lo	lo	hi			hi	lo	
	N=27	21	30	22	48	52	49	51	100
Minn. Math.Test score									
mean	34.00	32.00	33.87	31.23	33.13	32.75	32.76	33.10	32.93
S D	8.20	4.72	9.43	7.58	6.91	8.71	7.97	7.83	7.86
Predicted fall quarter grade point average									
mean	2.28	2.22	2.21	2.10	2.25	2.16	2.20	2.21	2.21
S D	.36	.25	.38	.38	.32	.38	.38	.33	.35
Obtained fall quarter grade point average									
mean	2.60	2.36	2.27	2.34	2.50	2.29	2.48	2.30	2.39
S D	.52	.54	.80	.74	.54	.77	.64	.70	.67
Grade point average discrepancy (algebraic)									
mean	.32	.14	.01	.23	.24	.13	.28	.01	.19
S D	.61	.42	.74	.63	.54	.70	.61	.63	.62
$r_{\text{MMS \& OFQGPA}}$	.16	.67	.35	.36	.33	.34	.28	.39	.33
[illegible]	.09	.64	.38	.52	.30	.43	.35	.44	.39

Table 3

Variance Indices and Mean Scores on each of Six RPM Tests and for the Total Measure Using Standard Scores for 100 University of Minnesota Institute of Technology Freshmen--Divided into Four Groups on the Basis of Relationships Between Predicted and Obtained Fall Quarter GPA.

Group	N	Aiming Mean	Aiming SD	Flexibility of Closure Mean	Flexibility of Closure SD	Number Facility Mean	Number Facility SD	Perceptual Speed Mean	Perceptual Speed SD	Speed of Closure Mean	Speed of Closure SD	Visualization Mean	Visualization SD	Total Measure Mean	Total Measure SD
								Variance							
High GPA correctly predicted	31	38.86	35.89	41.10	19.54	14.61	7.86	38.64	31.81	55.92	29.13	35.13	16.01	60.93	17.59
Low GPA correctly predicted	36	54.53	46.84	43.95	19.42	21.37	27.19	38.45	25.36	51.06	22.80	36.03	16.12	67.94	34.15
GPA under-predicted	18	38.48	14.58	40.09	21.02	18.86	13.47	38.99	32.87	59.13	19.94	38.19	22.12	64.88	20.56
GPA over-predicted	15	41.88	24.46	33.71	15.55	17.42	11.48	42.05	21.30	54.92	29.15	33.46	15.14	64.44	24.90
								Mean							
High GPA correctly predicted	31	49.28	7.43	51.88	6.36	50.85	8.32	51.35	7.56	50.39	7.43	51.08	7.11	50.80	5.44
Low GPA correctly predicted	36	50.13	6.91	50.30	9.25	49.38	8.44	49.91	7.05	50.13	6.08	50.79	7.17	50.10	5.28
GPA under-predicted	18	53.28	7.81	47.59	7.99	50.35	8.69	51.18	5.15	50.55	6.30	50.05	8.43	50.50	5.51
GPA over-predicted	15	47.20	8.20	48.26	5.75	49.28	12.72	45.98	11.81	48.20	8.57	45.78	10.86	47.45	8.37

Table 4

Comparison of Low Scoring-Low Variance, Low Scoring-High Variance, High Scoring-Low Variance, and High Scoring-High Variance Groups, Using the Aiming Test of the RPM[(1)], on Predicted and Obtained Fall Quarter Grade Point Average and Discrepancy Between Predicted and Obtained GPA.

Group	N	Predicted Fall GPA Mean	Predicted Fall GPA S D	Obtained Fall GPA Mean	Obtained Fall GPA S D	GPA discrepancy Mean	GPA discrepancy S D
Low mean low variance quadrant	30	2.26	.35	2.40	.69	.15	.60
Low mean high variance quadrant	18	2.20	.34	1.90	.68	-.30	.62
High mean low variance quadrant	18	2.35	.38	2.66	.56	.30	.60
High mean high variance quadrant	34	2.08	.31	2.50	.59	.42	.56

(1) Results reported here are based on raw scores of Aiming test. Corresponding table based on standard scores of Aiming test is almost identical.

Table 5

Correlations Between Predictors and Fall Quarter Grade Point Average for Groups with High and Low Variances on the RPM Number Facility Test.

$N^{(1)}$		Variance above median 47	Variance below median 46
Correlation between Math test and Fall quarter GPA		.34	.32
Correlation between HSR and Fall quarter GPA		.46	.48
Correlation between MSAT and Fall quarter GPA		.21	.15
Correlation between MSAT and HSR		.33	.26
Correlation between Math test and HSR		.26	.07
Math test	mean	31.62	33.07
	S D	5.88	8.53
HSR	mean	85.91	89.41
	S D	10.80	7.59
MSAT	mean	54.43	55.15
	S D	11.09	12.02
GPA	mean	2.41	2.36
	S D	.67	.70
Discrepancy between predicted and obtained GPA	mean	.23	.14
	S D	.64	.63

(1) HSR and MSA available for only 93 of 100 subjects.

Table 6

Relationships Between Variance and Predictability for 56 Subjects in Middle Ability Range.[1]

Correlation between discrepancy between predicted and obtained Fall quarter GPA and variance on Aiming test	r = -.02
Correlation between discrepancy between predicted and obtained Fall quarter GPA and variance on Number Facility test	r = .02
Correlation between above two variances	r = -.03
Correlation between predicted and obtained Fall quarter GPA	r = .51
Correlation between predicted and obtained GPA for 28 subjects with lowest variance on Aiming test	r = .58
Correlation between predicted and obtained GPA for 28 subjects with highest variance on Aiming test	r = .40
Correlation between predicted and obtained GPA for 28 subjects with lowest variance on Number Facility test	r = .58
Correlation between predicted and obtained GPA for 28 subjects with highest variance on Number Facility test	r = .46

(1) These subjects had mean scores on both Aiming and Number Facility tests that placed them within $\pm$ S.D. of the means for total group.

Table 7

Comparisons of Subjects Divided into Two Groups, Those Above and Those Below the Median Coefficient of Variation on the Number Facility Test.

(N=93[1])

	Low coefficient of variation group N=46		High coefficient of variation group N=47	
	Mean	S D	Mean	S D
Minnesota Math test	33.65	7.88	31.04	6.53
Predicted Fall quarter GPA	2.26	.37	2.14	.34
Obtained Fall quarter GPA	2.49	.53	2.28	.80
Difference between predicted and obtained	.23	.45	.13	.78
High school rank	88.89	8.88	86.43	9.94
MSAT	54.17	10.55	55.38	12.45
Correlated variables				
Minn. Math test and Fall quarter GPA	.46		.20	
Predicted and obtained Fall quarter GPA	.55		.26	
High school rank and Fall quarter GPA	.48		.41	
MSAT and Fall quarter GPA	.35		.10	
MMT and Winter quarter GPA	.34		.19	
High school rank and Winter GPA	.26		.52	
MSAT and Winter quarter GPA	.13		.13	
Fall quarter and Winter quarter GPA	.63		.73	
MSAT and HSR	.42		.22	
MMT and HSR	.34**		-.05	

[1]High School Rank and MSAT score available for only 93 subjects.

** $P < .05$